Quality checklist for decentralized evaluation

Evaluation report

Version June 2024

[Title of the decentralized evaluation]

**he evaluation firm (or team leader if hired as individual) should not submit the draft ER to WFP if it does not meet quality criteria. If the EM receives a draft that does not meet quality criteria, they should not submit to DE Quality Support (DEQS). REU DE Focal point (or CapQual incase of HQ DEs) as Quality Assurance Level 2 (QA2) should confirm that the ER meets the criteria and is ready for submission to DEQS. If it does not meet the criteria it should be returned to the firm (team leader). If the ER is submitted to DEQS without meeting the quality criteria, it will be returned which could result in delays in the evaluation process. If in doubt whether the ER is ready for submission to DEQS or the firm or the EM has reasons to want to submit incomplete ER to DEQAS, write to DE helpdesk wfp.decentralizedevaluation@wfp.org** [**wfp.decentralizedevaluation@wfp.org**](mailto:wfp.decentralizedevaluation@wfp.org)

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Overall** | | | | | |
| **General Criteria** | | | **Criteria met** | **Comments** | |
| **Length**   * Report does not exceed 30,000 words (or 34,500 for French or Spanish ERs), excluding executive summary; * Annexes do not exceed 40,000 words (or or 36,000 for French or Spanish ERs).   **Accessibility**   * Is the report is written in a clear and accessible language? * Does the structure of the findings allow readers to follow the argument and understand evaluation findings and conclusions? * Are key messages summarized and highlighted? * Does the report present a good balance between descriptive and analytical information, without excessible description and/or repetitions? * Are terminologies is used correctly including definitions where appropriate? * Is the report free from grammar, spelling or punctuation errors? * Are visuals, diagrams, charts and tables used in a clear and accessible way? * Is the language used in the report simple and clear, without jargon and excessively complex sentences and acronyms are used only where essential? * Is relevant information that can be found in other parts of the report is adequately signposted ?   **Credibility**   * Is the tone of the report constructive, balanced and sensitive to the political and cultural contexts? * Is the report objective and presents successes and failures, positive and negative sides in a balanced way * Is the language precise and professional, and appropriate for an official document, without room for interpretation? * Does the report raise critical issues when necessary and does not avoid or hide problems or issues? * Are the report’s findings and conclusions fully evidence-based? * Does the evaluation not draw conclusions unless evidence clearly points to them? | | |  |  | |
| **Editing** | | |  |  | |
| * Has the report applied WFP format and templates are applied? * Are all acronyms spelt out the first time they are used and listed? * Are paragraphs and pages numbered electronically? * Are sources provided for all data and quotes and footnotes used appropriately? * Is a table of contents included and lists tables, graphs, figures and annexes? * Are tables and diagrams used as relevant and are numbered? * Are annexes numbered and cross-referenced in the main document? | | |  |  | |
| **Cover page** | | |  |  | |
| * Has the cover page used the WFP DEQAS report template/layout (as adjusted for joint evaluations as appropriate)? * Is the title of the decentralized evaluation identical to that in the ToR (unless agreed otherwise)? * Is the date and status report (draft/final) of the indicated on the cover page? | | |  |  | |
| **Executive summary** | | | | | |
| **Expected content** | **Assessment criteria** | **Criteria Met** | | | **Comments/status** |
| The executive summary provides a complete and balanced synthesis of the evaluation findings, conclusions and recommendations and contains:   * An introduction, which details the main features of the evaluation including evaluation type, period being evaluated and commissioning office * Purpose and objectives, main users/intended audience, context * The subject of the evaluation, including geographical coverage, beneficiaries and resources raised against budget * Main features of the methodology (including rationale, sources, data collection and analysis methods used, major limitations) * Key findings on all the evaluation questions * Summarized conclusions and recommendations * If learning was a key focus of the evaluation, lessons learned identified are clearly summarized. | * Is the executive summary is coherent, self-explanatory and self-contained?can it be used as a standalone aid to decision-making? * Do the key findings in the summary provide answers to all evaluation questions? * Are conclusions , with clearly linked to findings presented without repeating the findings and/or introducing new information? * Are all recommendations presented in brief * Is the length 2,500 words or less (or 2,875 for French and Spanish ER) for CO-led DEs; and 3,000 words (or 3,450) for RB-led and HQ-led DEs? * If the evaluation TOR indicated learning objectives, does the ER include lessons i.e. key elements of learning derived from the evaluation findings and conclusions? |  | | |  |

|  |  |  |  |
| --- | --- | --- | --- |
| 1. **Introduction** | | | |
| **Expected content** | **Assessment criteria** | **Meets CRiteria** | **Comments** |
| **1.1 Evaluation features** | | | |
| This section provides a succinct overview of the evaluation features, providing clarity on why and how the evaluation was carried out and includes:   * The purpose / rationale for the evaluation, including why the evaluation is taking place at this point in time * Specific objectives of accountability and/or learning . * The scope of the evaluation, including any exclusions and reasons for them * Main stakeholders and users in the evaluationand how they will use the evaluation findings * A short presentation of the evaluation team, timing and duration of fieldwork | * Is the section succinct and focused without repeating information presented elsewhere in the ER? * Is the rationale of WHY the evaluation was commissioned clear? * Are the objectives of the evaluation clear? * Is it clear if or not more weight was given to a accountability or learning objective? , Is it explicitly justified in the report? * Is the scope of the evaluation clearly explained? * Is it clear from reading this section who the main stakeholders and users are? * Is the content sufficient to explain to readers how the evaluation was undertaken and to generate trust in the impartiality and credibility of the evaluation? * Are relevant annexes cross-referenced? |  |  |
| **1.2 Context** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| This section provides a succinct overview of the surrounding context directly relevant to the evaluation including:   * Poverty, food and nutrition security * Contextual aspects relevant to, and necessary to inform an understanding of the subject of the evaluation: geography, demography, including refugees and internally displaced persons where applicable; disasters and humanitarian protection where applicable * Government strategy, institutional capacities, policies and priorities, including normative instruments related to human rights, gender equality and equity and wider inclusion considerations * National indicators (education, health, nutrition, agriculture, gender inequality index) relevant to the evaluation subject, disaggregated by sex. * Key data and trends related to SDG 2/SDG 17 in the context (region, country subnational/local level) * Humanitarian issues, including migration patterns and host community/social tensions (IF APPLICABLE) * Gender, equity and wider inclusion dimensions of the context, including an intersectional analysis of specific social groups as appropriate * Key external events which led to any significant changes in WFP’s work (if applicable) * Features of international assistance in the area and related work of other key humanitarian/ development actors * Other WFP work in the area (if any) | * Is information is sufficient to understand the context within which the evaluand was implemented? * Are relevant indicators identified and latest trend data used or gaps mentioned where data does not exist? * Does the information draw from and is consistent with the inception report, updated where appropriate? * Is the information explicitly geared to the evaluation subject, rather than being generically presented? * Is there a good balance between amount of description of details and synthesis of what it all means? * Are authoritative or reliable sources used for important indicators? |  |  |
| **1.3. Subject being evaluated** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| This section provides a succinct overview of the evaluation subject, including factual information such as:   * Type of evaluation (activity, thematic area, transfer modality, pilot project) and subject of evaluation * Geographic scope of the evaluation subject * Relevant dates: approval date; start date; end date * Presentation of the logic of the intervention/theory of change/logical framework and its key assumptions, including intended results (planned outputs, outcomes and where appropriate, cross-cutting). If no ToC, the evaluation team should have recronstructed it at inception stage * Planned and actual outputs and activities * Beneficiary numbers (planned and actual) disaggregated by sex/age/activity * Amount of planned and actual transfers (tonnage/cash/vouchers), disaggregated by year and by activity (if multiple activities). For DEs that focus on an activity, transfer modality or pilot project, the modalities of intervention are presented and discussed (e.g. capacity strengthening, service delivery, etc.). * Main partners (government; NGOs; bilateral; multilateral, private sector) * Resources (over time, resources allocated to each activity, percentage funded of total requirements) and key donors. If subject funded from pooled funds or is a component of an activity, show resource allocated to the subject of the evaluation * Analytical basis of the subject (needs assessment, past evaluations/reviews, gender/social inclusion studies that informed the design of the subject). Any amendments to initial design and rationale. Includes a summary of findings from previous related evaluations * Other relevant preceding/ concurrent activities/ interventions * Maps/graphs for illustration | * Is it clear from reading this section on what the subject of evaluation is? * Is it clear when the subject was designed? * Does the ER highlights relevant evidence from past needs assessments, evaluations and reviews were relevant to the subject of evaluation and whether these analytical bases were used in the design? * Does the ER provided information on the budget i.e. USD value? * Does the ER provide summary of planned and actual outputs? (beneficiaries, food in metric tons, cash & voucher in USD as appropriate)? * Does the ER discuss the targeting/scope? * Does the ER discuss the planned and actual outcomes? * Is it clear who is involved in the implementation and extent of implementation? * Does the ER explain gender, equity and wider inclusion dimensions of the subject of the evaluation? * Are differences between original design and implementation explained if appropriate? * Does the ER describe he subject without pre-empting the analysis of the findings section? |  |  |
| **1.4. Evaluation methodology, limitations and ethical considerations** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| * Methodology used to answer evaluation * Evaluation criteria applied to the evaluation consistent with the purpose/scope of the evaluation and in relation to context. * The main evaluation questions in relation to each evaluation criterion * An evaluation matrix covering all evaluation questions * Gender, equity and wider inclusion considerations across the evaluation questions and data collection/analysis methods iwhere relevant * Specific data collection and analytical methods, data sources and sampling used (area and population, rationale for selection, representation) * Assessment of the strength of the intervention’s monitoring processes and data in relation to gender, equity and wider inclusion considerations are provided * Approaches to triangulation with a description on how validity and reliability was addressed * Limitations of methods or data availability encountered, and mitigation strategies applied/implications for the findings * How the evaluation has considered and managed ethics, and what safeguards were in put in place during the evaluation are detailed. * If relevant, how the analysis has been conducted against international humanitarian principles | * Does the information provided enable reader to understand how the design of the evaluation and how it was conducted? * Is the methodological approach comprehensive and systematic in ways that generate trust in the credibility of the evaluation? * Is the methodological approach coherent, logical and in line with the ToR and inception report? * Does the evaluation matrix identify context-specific sub-questions aligned with the main evaluation questions in the inception report, as well as indicators, data sources, data collection methods and analysis methods for each sub-question? * Do the methods identified allow for analysis of GEWE, inclusion and equity dimensions? * Are specific methods and sampling relevant to the methodological approach and feasible/appropriate? * Are analytical plans and methods described in ways that assure confidence in the evidence base (how data collected have been processed, analysed and interpreted)? * Does the ER highlight limitations, including on the validity and reliability of primary data collected, and their effects on the evidence base and how they were addressed or implications if they were not? * Does the ER explains how any risks to implementation of the methodology were addressed? * As ethical safeguards and appropriate mitigation measures that were taken explained? * If any risks and ethical issues anticipated in ToR/IR explicitly and specifically addressed in the ER? * Evaluation deliverables do not include any personal identifiable information * Is the full methodology provided in Annexes and cross-referenced in the main report? * Was the methodology implemented as planned in the IR? If not, does the ER reference any aspects of the methodology that were not implemented as planned? |  |  |
| 1. **Evaluation findings** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| * Findings structured around the evaluation questions and corresponding subquestions as presented in the evaluation matrix in inception report * Any gap in addressing evaluation questions and sub-questions as compared to the scope agreed in the IR * Limitations or gaps in evidence or caveats based on contextual or other concerns including where evidence is inconclusive * Findings assess WFP's contributions to results (or lack of) while taking into consideration the implementation context and other effects produced. * Positive or negative unanticipated effects. If none were found, the ER indicates that explicitly . | * Do the findings explicitly address all the evaluation questions as per the evaluation TOR? * If ther are gaps or omissions, are they explained and justified? * Are the findings clearly aligned with the evaluation matrix? * Are findings transparently generated on the basis of triangulated evidence presented in the report?. * Is evidence accurate, consistent and reliable; withsources provided for data/quotes?. * Is the ER devoid of findings is based on unsubstantiated opinions or attributions to ‘one person’? * Do the findings take into consideration the perspectives from different stakeholder groups including women and young people, and carefully consider gender, equity and wider inclusion dimensions, including through the analysis of data disaggregated by sex, age, disability, etc.? * Are findings presented in a culturally sensitive manner that respects confidentiality, protection of sources and dignity of those interviewed? * Are limitation or gaps in the evidence base clearly explained, and any places where the evidence is inconclusive clearly be explained? * Are findings balanced, presenting both successes and failures, positive and negative aspects in a constructive manner? * Are findings adequate to substantiate the conclusions and recommendations made by the evaluation team? * Do findings consider how recommendations from previous evaluations have been addressed or not addressed (where such evaluations exist)? * Do findings assess performance against international humanitarian principles (if applicable)? * Are visual aids used in a clear an accessible way to illustrate the arguments? * Do the findings balance descriptions and synthesis? * Do the findings include positive or negative unitended results (or indication that none were found)? * Do sources of quotations of interviews, focus group discussions and events not allow for personal identification? |  |  |
| 1. **Conclusions and recommendations** | | | |
| **3.1 Conclusions** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| * ABalanced assessment of both positive and negative, intended and unintended results as presented in the evaluation findings. * Conclusions may be organized by evaluation question, evaluation criteria or by theme * The conclusions synthesize the findings up to a higher level, connecting findings across different evaluation criteria and responding to the “so what” question (rather than simply summarizing the evidence) * The conclusions consider gender equality and the empowerment of women (GEWE), equity and wider inclusion dimensions * Conclusions address expliclity positive and/or negative unintended results | * Do conclusions follow logically from the findings, with clear and explicit linkage to them? * Are conclusions succinct,presenting a conclusive picture based on findings rather than repeating the findings? * Are conclusions devoid of new evidence that has not been presented in the findings? * Are conclusions are impartial/based on evidence already presented and free of bias or evaluator’s opinion? * Do conclusions reflect both strengths and areas for improvements? * Do conclusions include recognition of different stakeholder groups, gender and equity aspects and how the findings/evidence relate to them? |  |  |
| **3.2 Lessons (Required if the evaluation TOR had a learning objective)** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| * Lessons derived from conclusions that are worth noting that do not lend themselves to concrete recommendations but rather reflect things that are considered good practices, things that have not worked, reasons why things have worked etc * Lessons presented in a way that can contribute to wider organizational learning for WFP and partners beyond the commissioning office. (this implies clearly identifying the conditions/ situation(s) for which they are valid) * They can be positive or negative | * If the evaluation TOR indicated a learning focus, does the ER include a lessons learnt section? * Are lessons clearly extracted from on the findings and conclusions of the evaluation? * Do the lessons contribute to wider organizational learning in WFP and/or can guide future action? * Do the lessons have the potential for wider application and use beyond the context of the evaluation? |  |  |
| **3.3 Recommendations** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| * A short paragraph describing each recommendation that is expliclity derived from the evaluation conclusions * Recommendations presented as per the table provided in the evaluation report template | * Is each recommendation relevant and well aligned to the evaluation purpose and objectives * Does each recommendation address critical areas identified by findings with no gaps? * Do recommendations clearly and logically derive from findings and conclusions? * Are recommendations internally consistent and take account of interdependencies between them? * Are recommendatons objective and constructive? * Are recommendations sufficiently specific to be actionable and provide a clear direction of intended change, while at the same time leaving room for implementers and users to fine-tune their implementation approach? * Do recommendations focus on what actions WFP can take (or other commissioners in the case of joint evaluations)? * Are recommendations oriented towards actions that clearly deliver benefits in proportion to their potential costs of implementation? * Do recommendations address GEWE issues (if findings and conclusions call for these) and provide priorities for action to improve GEWE and reflect, where appropriate, equity and inclusion dimensions? * Are recommendations clearly targeted making clear/specifying who should implement them? * Are recommendations prioritized, distinguishing between those that are high or medium priority?   Are recommendations concise?   * Are recommendations expressed in active language indicating what SHOULD be done? * Are recommendations grouped /sequenced based on criteria (e.g. strategic/ operational)? * Is there a propose clear timeframe for action for each recommendation? * Do recommendations adhere to do-no-harm principle? |  |  |
| **Annexes** | | | |
| **Expected content** | **Assessment criteria** | **Meets Citeria** | **Comments** |
| Annexes support and expand on text in the main report and do not include all working documents.  Mandatory annexes include:   * Summary ToR [Ask WFP to provide] * Evaluation timeline * Methodology with intended and actual methods applied and any limitations to validity. Where appropriate, it should provide reflection on experience and lessons for future evaluation. It should indicate the extent to which GEWE considerations were incorporated where applicable (e.g. GEWE as a cross-cutting theme), and how GEWE was integrated into data collection methods) * Evaluation matrix * Data collection tools * Fieldwork agenda * Findings – conclusions - recommendations mapping * Key informants overview * Bibliography * Acronyms | * Annexes do not exceed 40,000 words * Are mandatory annexes are all included? * Are additional technical annexes are relevant and necessary to supplement the main text? * Are all annexes are listed, numbered and referenced where appropriate in the main report? * Is the annex "Key informants’ overview” devoid of any personal identifiable information and instead include only affiliation and/or indications on municipalities? |  |  |